Reference amounts utilised in front of package nutrition labelling; impact on product healthfulness evaluations

M. M. Raats, S. Hieke, C. Jola, C. Hodgkins, J. Kennedy and J. Wills

Raats, M.M., et al. 2015. Reference amounts utilised in front of package nutrition labelling; impact on product healthfulness evaluations. *European Journal of Clinical Nutrition*. 69(5): pp.619-625. doi: 10.1038/ejcn.2014.190

The published article is available from doi: 10.1038/ejcn.2014.190

Reference amounts utilised in front of package nutrition labelling; impact on product healthfulness evaluations

MM Raats[1] , S Hieke[2] , C Jola[3] , C Hodgkins[1] , J Kennedy[2,4,5] and J Wills[2]

[1] Food, Consumer Behaviour and Health Research Centre, University of Surrey, Guildford, UK; [2] European Food Information Council Brussels, Belgium; [3] School of Social and Health Sciences-Psychology, Abertay University Dundee, UK; [4] Alltech, Dunboyne, Ireland and [5] UCD School of Public Health, Physiotherapy and Population Science, University College Dublin, Ireland.

Correspondence: Professor MM Raats, Food, Consumer Behaviour and Health Research Centre, School of Psychology, Faculty of Arts and Human Sciences, University of Surrey, Guildford, GU2 7XH, UK. E-mail: m.raats@surrey.ac.uk

ABSTRACT

BACKGROUND/OBJECTIVES: The research question addressed in this paper is how different reference amounts utilised in front of package nutrition labelling influence evaluation of product healthfulness. **SUBJECTS/METHODS:** A total of 13 117 participants from six European countries (Germany, UK, Spain, France, Poland and Sweden) were recruited via online panels. A mixed between/within-subject factorial design was employed with food (biscuits, sandwiches, yogurts), healthfulness and presence of Guideline Daily Amounts as within-subjects factors and reference amount ('per 100 g', 'typical portion', 'half portion') and country as between-subjects factors. **RESULTS:** Overall, people correctly ranked foods according to their objective healthfulness as defined by risk nutrients alone, and could distinguish between more and less healthful variants of foods. General healthfulness associations with the three product categories do not appear to have had a strong influence on product ratings. This study shows that where the reference amount of 'per 100 g' is very different from the 'typical' portion size, as was the case for biscuits, products with a 'per 100 g' label are rated significantly less healthful than the 'typical' or 'half typical' portions. **CONCLUSION:** The results indicate that across the three food categories, consumers do factor the reference amount, that is, the quantity of food for which the nutritional information is being presented, into their judgements of healthfulness. Therefore, appropriate reference amounts are also of importance for the effective presentation of nutritional information.

INTRODUCTION

On both back of package and front of package (FoP), the reference amount chosen forms an integral part of the nutrition information displayed on packaging. As the provision of nutrition information is regarded by policy makers as a means of encouraging consumers towards more healthful food choices, it is important to better understand the potential impact of reference amounts on healthfulness evaluations. Reference amounts utilised in nutrition labelling are predominantly 'per 100 g/100 ml', 'per unit' and 'per portion'.

Different nutrition labelling formats have been shown to influence food product healthfulness evaluations,[1–5] however, the experimental design and focus of analysis has generally been on the energy and nutrition content of the labels. So far, the role of reference amounts, an intrinsic element of many of the FoP labelling schemes, appears not to have been systematically evaluated.

This is despite earlier research on back of package nutrition information provision suggesting that the way in which reference amount information is presented can have a substantial effect on food choices[6] and that consumers find it difficult to compare products when the nutrition information is presented using different reference amounts.[7,8]

Presentation of nutrition information on food packaging in the EU is governed by the recently adopted regulation on Food Information to Consumers,[9] which requires a mandatory nutrition declaration on pre-packed foods expressed 'per 100 g' or 'per 100 ml' on back of package; however, this can also be supplemented by nutrition information 'per portion'. Furthermore, in the EU, if a manufacturer chooses to repeat nutrition information on FoP, they can express the four key nutrients: fat, saturated fats, sugar and salt in terms of portion only, although the energy must be expressed as both 'per 100 g/100 ml' and 'per portion'. Where nutrition information per portion is provided, the portion on which the nutrition values are based must be quantified in close proximity to the nutrition information. Therefore, because of an increased occurrence of FoP nutrition labelling, understanding the impact of reference amounts on health inferences is becoming increasingly important.

In their 2005 review of the nutrition labelling literature, Cowburn and Stockley[10] concluded that consumers found portion information difficult to interpret. Research since 2005 has shown that consumer confusion related to portion information is still an issue.[11,12] For example, when comparing eight executions of FoP calorie flags, Van Kleef et al.[5] reported that consumers found the 'calories per serving flag' option to be most preferable but only under the condition that the portion was perceived to represent a realistic and easy-to-understand consumption unit. There is on-going debate as to whether nutrition information is best presented 'per 100 g/100 ml' or 'per portion' [8,13,14] particularly amongst those concerned with the development of nutrient profiling models which attempt to define 'healthy' versus 'unhealthy' foods. Many consider that 100 g/100 ml best serves as a comparator between foods that are often presented in a range of different portion sizes, whereas others argue that a food containing high levels of an 'unhealthy' nutrient 'per 100 g' may only make a small negative contribution to a person's diet if the 'typical' portion for that food is much smaller than 100 g.[13]

Product category associations (e.g., nutrition value of chocolate versus yogurt) have been shown to be one of the cues used by consumers to form judgements about a product's healthfulness.[15,16] On the basis of the evaluability principle, it is also argued that 'numerical [nutrient content] information lacks meaning by itself and has to be compared with other information to be interpreted meaningfully'. [17] In other words, a comparison baseline, for example a Guideline Daily Amount (GDA), may be essential for the processing of numerical nutrition information.[15,18] Foods are often categorised as 'healthy' or 'diet'; and it has been shown that foods considered part of a 'healthy' category can elicit a systematic underestimation of their energy content, resulting in higher intake quantities regardless of the portion size.[16,19]

This study explores to what extent participants infer healthfulness of foods across three very different food categories, in three reference amounts, using an FoP labelling scheme in which nutrition information is presented with and without percentage GDAs. GDA schemes express values for energy, sugar, fats, saturated fats and salt that a portion of the food contains as a percentage contribution to the daily requirements of an average reference adult. GDAs were derived from the COMA report[20] on Dietary Reference Values and are a labelling scheme used widely across Europe.[21]

MATERIALS AND METHODS

Data on participants' perceptions of relative, as well as within-category, healthfulness of foods were collected, using a mixed between/within-subject factorial design. Participants were randomly assigned to one of three reference amount conditions ('per 100 g', 'typical' portion, 50% reduction on 'typical' portion) and rated three different foods (biscuits (i.e. cookies in North America), sandwiches, yogurts) on their healthfulness. For each food, participants rated two pairs of products within the same category, with one product more healthful than the other. Participants rated products without and then with GDA energy/nutrient labelling.

Basing the task on a direct comparison of two products but at the same time asking respondents to evaluate healthfulness within a product category was expected to yield more insights on how health inferences are made on food products and the role of reference amounts.

Recruitment of participants

A quota sample of 13 117 participants was obtained from online research panels in six European countries (Germany n = 2171, UK n = 2155, Spain n = 2206, France n = 2209, Poland n = 2169 and Sweden n = 2207; see Table 1). Quotas were also applied for age, gender and level of education. Body mass index was calculated based on participants' self-reported height and weight, as weight (in kg) divided by height (in m).2 Participants were classified according to social grade using the National Readership Survey (NRS) classification system[22] based on questions on the household's chief income earner (namely working status, occupation and number of people responsible for).

Materials

Before the online experiment started, participants were presented with an example of a label using GDA and a short definition: 'Guideline Daily Amounts (GDAs) are guidelines about the amount of some nutrients e.g. fat, salt, that a person should be eating in a day.' Within the experiment, participants were randomly assigned to one of three FoP reference amount groups, each of which was presented with nutrition information for either per 100 g, a 'typical' portion size or half of a 'typical' portion size, henceforth referred to 'per 100 g', 'typical portion' and 'half portion', respectively. The product nutrition information is shown in Table 2.

When selecting the food categories to include, it was necessary to consider the different food cultures in the participating countries and select categories that were familiar in all four. Biscuits, sandwiches and yogurts were chosen as they satisfied this primary criterion and they represented a wide range of portion sizes; biscuits being typically a snack food that tends to be presented in small portions or units, whereas for a product that forms the main component of a meal, such as sandwiches, portions are larger, with yogurts falling somewhere in between. Following a review of the 'typical' portion sizes on the market for each of the three chosen food categories, a standardised 'typical portion' was set for each of the three food categories; biscuits 18 g, sandwiches 250 g and yogurts 150 g. The second portion condition tested was then set as a 50% reduction on this standardised 'typical portion' condition to see if health inferences were impacted by a reduction in portion size. The 'per 100 g' label was included as a comparator between the foods.

To facilitate the final food stimuli selection within each category, it was necessary to map the relative healthiness of the foods both within and across the food categories. This was achieved by

application of the SSAg/1 nutrient profiling system, one of the approaches considered in the work to support the UK Food Standards Agency initiatives to address food advertising.[23] SSAg/1 scores start at zero for the most healthful foods and increase in units of 1 per 10% increase in GDA of the energy and each nutrient contained in 100 g of the food. SSAg/1 was considered to be the most appropriate objective health scoring model for this study because it results in an absolute score for each food based only on energy and the main risk nutrients, that is, saturated fat, sugar and salt, without taking into consideration any positive aspects of the food, such as levels of micronutrients or fibre, and this maps onto the communication elements typically presented in the majority of FoP labelling. SSAg/1 scores for each of the foods are detailed in Table 2. The final two food variants representing different levels of healthfulness within each category were selected by reviewing the nutrition values of real foods on the market and selecting those that represented a realistic upper, lower and mid-range within each category.

Procedure

Participants were simultaneously presented with two FoP labels and asked to rate the healthfulness for both products individually by positioning the products in question on a slider with the mouse to a location between 'least healthy food you can think of' on one end (left) to 'most healthy food you can think of' on the other end (right) presented below the FoP labels. Each label (more healthful and less healthful version) was randomly assigned to the left or the right side of the screen. Participants rated both labels (more healthful, less healthful) first without GDA and then with GDA (the first set of labels showed no GDA information, only absolute values for energy and nutrients, whereas the second set showed GDA information on top of the absolute values, in per cent). This was repeated for each of the three types of foods presented in random order. However, the label containing GDA information was always presented after the one with gram information only. Hence, each participant provided 12 healthfulness ratings on six presentation slides that were either based on 'typical portion', 'half typical portion' or 'per 100 g', dependent on their assigned group. The slider for the healthfulness ratings was continuous and measured 101 positions from zero for the least healthful imaginary product to 100 for the healthiest imaginary product. Using this qualitative approach to healthfulness evaluation, a limitation of existing literature could be overcome: many studies use calorie estimates as a proxy for judging the healthfulness of a product.[19] This method is limited as calories play only one part in the overall healthfulness of a product or even an entire diet. For the purpose of comparison to the 'objective' SSAg/1 scores, the healthfulness scores from participants were reversed and scaled from 0 to 15.

Analysis

A 3 food (biscuits, sandwiches, yogurts) × 2 healthfulness (more healthful, less healthful) × 2 FoP labelling conditions (Baseline versus FoP label) × 3 reference amount ('per 100 g', 'typical portion', 'half portion') × 6 country mixed measures analysis of variance was performed in SPSS version 19.0 (SPSS, Chicago, IL, USA), using the participants' perceived healthfulness ratings as the dependent variable. The Greenhouse-Geisser method, which corrects for any violation of the sphericity assumption, was used to calculate F and P. Measures of effect size are reported using partial eta-squared (ηp^2), which represents the strength of the association of each independent variable (FoP condition, reference amount, country) with the dependent variable (participant's perceived healthfulness ratings), after the effects of all other independent variables were accounted for.

Table 1. Participant characteristics (as percentages of the total samples)

		France *n* 2209	Germany *n* 2171	Poland *n* 2169	Spain *n* 2206	Sweden *n* 2207	UK *n* 2155	Total *n* 13117
Gender	Male	45.5	47.0	42.4	45.9	38.2	47.0	44.3
	Female	54.5	53.0	57.6	54.1	61.8	53.0	55.7
Age	18-29	24.8	22.2	32.1	27.5	21.9	20.0	24.7
	30-39	21.9	20.9	24.5	30.8	25.4	24.4	24.7
	40-49	23.2	28.1	22.0	25.9	25.0	25.8	25.0
	50-64	30.1	28.9	21.4	15.9	27.6	29.7	25.6
Education	None	1.6	0.5	0.1	0.4	0.2	0.7	0.6
	Primary school	1.7	13.4	2.7	3.8	6.5	.1	4.7
	Secondary school to age 15/16	20.1	35.4	3.1	11.7	3.0	28.4	16.9
	Secondary school to age 17/18	47.2	21.7	64.8	43.7	49.3	48.5	45.9
	College/Undergraduate	15.5	8.2	8.3	20.4	20.5	14.3	14.6
	University/Post graduate	14.0	20.9	20.9	20.1	20.4	8.0	17.4
Social grade	A: Higher managerial, administrative or professional	2.4	2.8	7.3	4.6	3.4	5.7	4.3
	B: Intermediate managerial, administrative or professional	13.3	15.1	11.7	15.8	16.4	15.5	14.6
	C1: Supervisory or clerical and junior managerial, administrative or professional	29.7	29.9	34.5	27.5	28.8	30.2	30.1
	C2: Skilled manual workers	41.4	35.5	35.3	37.0	38.8	32.7	36.8
	D: Semi and unskilled manual workers	3.3	3.4	3.8	5.1	3.6	3.9	3.8
	E: Casual or lowest grade workers, pensioners, and others who depend on the welfare state for their income	10.0	13.4	7.4	10.0	9.0	12.1	10.3
Body Mass Index (BMI)[1]	Underweight (BMI≤18)	4.6	2.7	3.5	2.9	2.5	3.1	3.2
	Normal weight (18<BMI<25)	52.4	46.8	49.0	50.5	48.0	42.2	48.2
	Overweight (25≤BMI<30)	30.3	32.8	32.0	34.3	32.4	32.1	32.3
	Obese (BMI(BMI>30)	12.7	17.7	15.5	12.3	17.1	22.7	16.3

Table 2. Nutritional profile of food stimuli and label details Food Level of healthfulness

Food	Level of healthfulness	SSAg/1 score	Reference amount[1]	Weight stated on label	Amount of energy and nutrients					%GDAs				
					energy (kcals)	sugar (g)	fat (g)	saturated fats (g)	salt (g)	energy (kcals)	sugar	fat	saturated fats	salt
										2000	90	70	20	6
biscuit	more healthful	5	100 g	100	425	18.7	9.1	3.5	0.8	21	21	13	18	13
			half typical portion	9	38	1.7	0.8	0.3	0.1	2	2	1	2	1
			typical portion	18	77	3.4	1.6	0.6	0.1	4	4	2	3	2
	less healthful	12	100 g	100	535	25.0	34.0	20.0	0.3	27	28	49	100	5
			half typical portion	9	48	2.3	3.1	1.8	0.0	2	3	4	9	0
			typical portion	18	96	4.5	6.1	3.6	0.1	5	5	9	18	1
sandwich	more healthful	0	per 100 g	100	160	1.2	3.7	0.9	0.5	8	1	5	5	8
			half typical portion	125	200	1.5	4.6	1.1	0.6	10	2	7	6	10
			typical portion	250	400	3.0	9.3	2.3	1.3	20	3	13	11	21
	less healthful	4	100 g	100	275	2.6	13.6	5.6	1.0	14	3	19	28	17
			half typical portion	125	344	3.3	17.0	7.0	1.3	17	4	24	35	21
			typical portion	250	688	6.5	34.0	14.0	2.5	34	7	49	70	42
yogurt	more healthful	0	100 g	100	70	7.8	1.5	0.9	0.2	4	9	2	5	3
			half typical portion	75	53	5.9	1.1	0.7	0.2	3	7	2	3	3
			typical portion	150	105	11.7	2.3	1.4	0.3	5	13	3	7	5
	less healthful	3	100 g	100	159	10.7	11.3	8.0	0.1	8	12	16	40	2
			half typical portion	75	119	8.0	8.5	6.0	0.1	6	9	12	30	1
			typical portion	150	239	16.1	17.0	12.0	0.2	12	18	24	60	3

[1] not stated on the label

RESULTS

A mixed-design analysis of variance with food, healthfulness and presence of GDA as within-subjects factors and reference amount ('per 100 g', 'typical portion', 'half portion') and country as between-subjects factors (Table 3) revealed a main effect of healthfulness F(1, 13 099) = 18 537.2, P = 0.000, ηp^2 = 0.586. The least healthful products (based on the SSAg/1 score) were rated significantly less healthful (mean = 10.24) than more healthful products (mean = 5.20). Participants could clearly distinguish between the healthfulness of more versus less healthful products. The main effect for healthfulness was qualified by interactions between food and healthfulness, F1 (1.997, 26 162.8) = 365.8, P = 0.000, ηp^2 = 0.027; and healthfulness and country, F(5, 13 099) = 68.0, P = 0.000, ηp^2 = 0.025. Across the product categories, yogurts and sandwiches were evaluated as less healthful than their objective scores, whereas the healthfulness of biscuits was overestimated. In addition, ratings for the more and less healthful variant were very similar across all three categories (Figure 1), with means ranging from 9.55 to 10.99 for the less healthful and from 4.55 to 6.44 for the more healthful. This demonstrates that respondents used the full range of the available scale for each category and that product evaluations were carried out within, rather than across the three food categories.

There were also smaller, but significant, main effects for food (F1 (1.9, 25 248.8) = 710.5, P = 0.000, ηp^2 = 0.051) and for country, F(5, 13 099) = 56.9, P = 0.000, ηp^2 = 0.021 (Table 3). Respondents in Sweden (mean = 8.07) rated products as least healthful, whereas those in Poland (mean = 7.45) and France (mean = 7.49) rated products as significantly more healthful than those in the other countries. Germany (mean = 7.75), Spain (mean = 7.76) and the UK (mean = 7.78) did not differ from each other but were different from both the other two groups (Table 4). However, overall, biscuits were perceived as least healthful (mean = 7.96) closely followed by sandwiches (mean = 7.84) and yogurts were perceived as most healthful (mean = 7.36). The small but significant interaction between food and country, F1 (9.6, 25 248.8) = 27.1, P = 0.000, ηp^2 = 0.010, indicates that the healthfulness ratings of foods varied according to country (Table 3). The significant interaction between healthfulness and country, F(5, 13 099) = 68.0, P = 0.000, ηp^2 = 0.025, indicates that the healthfulness ratings of the more and less healthful variants of foods varied according to country (Table 3). The mean scores for more and less healthful foods ranged from 4.19 in Poland to 6.31 in France, respectively.

A small but significant main effect for reference amount was observed, F(2, 13 099) = 72.3, P = 0.000, ηp^2 = 0.011 (Table 3). Post hoc analysis revealed that participants perceived 'half portions' (mean = 7.51) to be significantly more healthful than 'typical portions' (mean = 7.85) or per 100 g (mean = 7.79) (Table 4). The significant interaction between food and reference amount, F1 (3.9, 25 248.8) = 522.6, P = 0.000, ηp^2 = 0.074 indicating that the healthfulness rating variation according to reference amount was differential across the foods (Greenhouse-Geisser correction utilised to correct for the violation of the sphericity assumption) (Table 3). However, on the whole, across all three foods, the larger the reference amount the less healthful it was perceived to be.

Table 3. Between- and within-subjects effects for healthfulness ratings of all foods across all countries

Interactions	F	df	Sig.	ηp²
Tests of Between-Subjects Effects				
Reference amount	**72.29**	**2, 13099**	**0.000**	**0.011**
Country	**56.91**	**5, 13099**	**0.000**	**0.021**
Reference amount*Country	**2.33**	**10, 13099**	**0.010**	**0.002**
Tests of Within-Subjects Effects				
Healthfulness	**18537.22**	**1, 13099**	**0.000**	**0.586**
Healthfulness*Reference amount	**9.05**	**2, 13099**	**0.000**	**0.001**
Healthfulness*Country	**67.98**	**5, 13099**	**0.000**	**0.025**
Healthfulness*Reference amount*Country	**3.58**	**10, 13099**	**0.000**	**0.003**
Food	**710.49**	**1.9, 25248.8**	**0.000**	**0.051**
Food*Reference amount	**522.60**	**3.9, 25248.8**	**0.000**	**0.074**
Food*Country	**27.07**	**9.6, 25248.8**	**0.000**	**0.010**
Food*Reference amount*Country	**14.38**	**19.3, 25248.8**	**0.000**	**0.011**
Food*Healthfulness	**365.76**	**2.0, 26162.8**	**0.000**	**0.027**
Food*Healthfulness*Reference amount	**11.36**	**4.0, 26162.8**	**0.000**	**0.002**
Food*Healthfulness*Country	**19.93**	**10.0, 26162.8**	**0.000**	**0.008**
Food*Healthfulness*Reference amount*Country	**6.74**	**20.0, 26162.8**	**0.000**	**0.005**
GDA	**4.50**	**1, 13099**	**0.034**	**0.000**
GDA*Reference amount	**4.80**	**2, 13099**	**0.008**	**0.001**
GDA*Country	**13.31**	**5, 13099**	**0.000**	**0.005**
GDA*Reference amount*Country	**5.29**	**10, 13099**	**0.000**	**0.004**
GDA*Food	**10.87**	**2.0, 26142.3**	**0.000**	**0.001**
GDA*Food*Reference amount	**32.74**	**4.0, 26142.3**	**0.000**	**0.005**
GDA*Food*Country	**8.56**	**10.0, 26142.3**	**0.000**	**0.003**
GDA*Food*Reference amount*Country	**6.12**	**20.0, 26142.3**	**0.000**	**0.005**
GDA*Healthfulness	**100.01**	**1, 13099**	**0.000**	**0.008**
GDA*Healthfulness*Reference amount	2.65	2, 13099	0.071	0.000
GDA*Healthfulness*Country	1.75	5, 13099	0.119	0.001
GDA*Healthfulness*Reference amount*Country	**9.27**	**10, 13099**	**0.000**	**0.007**
GDA*Food*Healthfulness	**21.71**	**2.0, 26155.6**	**0.000**	**.0002**
GDA*Food*Healthfulness*Reference amount	**5.27**	**4.0, 26155.6**	**0.000**	**0.001**
GDA*Food*Healthfulness*Country	**6.92**	**10.0, 26155.6**	**0.000**	**0.003**
GDA*Food*Healthfulness*Reference amount*Country	**6.48**	**20.0, 26155.6**	**0.000**	**0.005**

Bold print indicates significant results at $p \leq 0.05$. Greenhouse-Geisser corrections were utilised to correct for the violation of the sphericity assumption where appropriate

Despite the significant main effect observed for GDA, $F(1, 13\,099) = 4.50$, $P = 0.034$, $\eta p^2 = 0.000$ (Table 3), the effect size demonstrates that inclusion of percentage GDAs on the label had little effect on perceived healthfulness ratings over and above the provision of absolute values for energy (kcal) and nutrients (g).

When separate analyses of variances for each food were conducted (i.e., with healthfulness and presence of GDA as within-subjects factors and reference amount ('typical portion', 'half portion', 'per 100 g') and country as between-subjects factors), the significant main effect for reference amount was maintained (Table 5). However, the effect size increased and was greater for biscuits $F(2, 13\,099) = 425.44$, $P = 0.000$, $\eta p^2 = 0.061$ than for sandwiches $F(2, 13\,099) = 185.22$, $P = 0.000$, $\eta p^2 = 0.028$) or yogurts $F(2, 13\,099) = 70.37$, $P = 0.000$, $\eta p^2 = 0.011$) (Table 5). For biscuits, participants rated the product with a 'per 100 g' label as significantly less healthful, than the 'typical' or 'half typical' reference amounts presented as 18 g and 9 g, respectively. For sandwiches, the 'per 100 g' label was rated significantly more healthful than the 'half typical' (125 g) or 'typical' (250 g) reference amounts. For yogurts, where the 'per 100 g' label was closer to the reference amounts shown (i.e., 75 g and 150 g), the effect size was much smaller (Figure 1) with no significant difference between the healthfulness ratings for the 'half typical' (75 g) reference amount and the 'per 100 g' label. The 'typical' (150 g) portion of yogurt was, however, rated as significantly less healthful than the product with the 'per 100 g' label and the 'half typical' reference amount.

Table 4. Mean and standard error of the mean rescaled healthfulness ratings for all foods

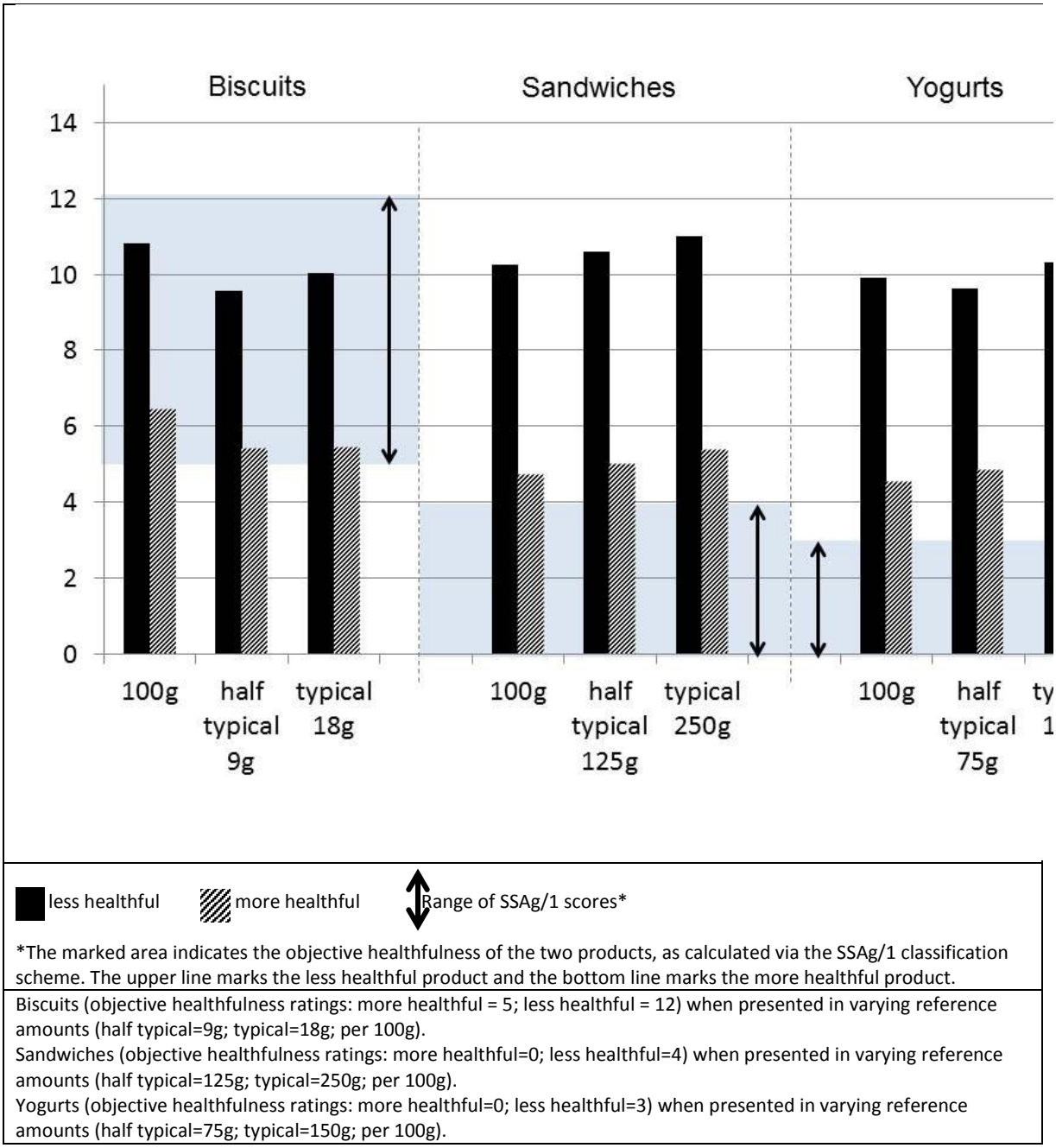
		All foods		Biscuits		Sandwiches		Yogurts	
		Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM
Portion[1]	Per 100 g	7.79[b]	0.02	8.64[c]	0.03	7.49[c]	0.03	7.23[a]	0.03
	Half typical	7.51[a]	0.02	7.48[a]	0.03	7.81[a]	0.03	7.24[a]	0.03
	Typical	7.84[b]	0.02	7.75[b]	0.03	8.20[b]	0.03	7.61[b]	0.03
Country[1]	France	7.49[a]	0.03	7.50[a]	0.04	7.62[a]	0.04	7.34[a]	0.04
	Germany	7.75[b]	0.03	7.96[bc]	0.04	7.90[bc]	0.04	7.39[a]	0.04
	Poland	7.46[a]	0.03	7.84[b]	0.04	7.64[b]	0.04	6.89[b]	0.04
	Spain	7.76[b]	0.03	7.92[b]	0.04	7.96[b]	0.04	7.39[a]	0.04
	Sweden	8.07[c]	0.03	8.42[d]	0.04	8.03[d]	0.04	7.76[c]	0.04
	UK	7.78[b]	0.03	8.09[c]	0.04	7.86[c]	0.04	7.38[a]	0.04

[1]Within a food, means with the same superscripts do not significantly differ from each other.

Table 5. Between- and within-subjects effects for healthfulness ratings of biscuits, sandwiches and yogurts for all countries

	Biscuits				Sandwiches				Yogurts			
	F	df	Sig.	ηp^2	F	df	Sig.	ηp^2	F	df	Sig.	ηp^2
Tests of Between-Subjects Effects												
Reference amount	**425.44**	**2**	**0.000**	**0.061**	**185.224**	**2**	**0.000**	**0.028**	**70.37**	**2**	**0.000**	**0.011**
Country	**54.05**	**5**	**0.000**	**0.020**	**21.197**	**5**	**0.000**	**0.008**	**58.55**	**5**	**0.000**	**0.022**
Reference amount*Country	**8.44**	**10**	**0.000**	**0.006**	**2.192**	**10**	**0.016**	**0.002**	**9.92**	**10**	**0.000**	**0.008**
Tests of Within-Subjects Effects												
Healthfulness	**9611.98**	**1**	**0.000**	**0.423**	**15193.6**	**1**	**0.000**	**0.537**	**12851.32**	**1**	**0.000**	**0.495**
Healthfulness*Reference amount	**8.76**	**2**	**0.000**	**0.001**	.24	2	0.790	0.000	**20.12**	**2**	**0.000**	**0.003**
Healthfulness*Country	**48.12**	**5**	**0.000**	**0.018**	**36.35**	**5**	**0.000**	**0.014**	**71.57**	**5**	**0.000**	**0.027**
Healthfulness*Reference amount*Country	**1.86**	**10**	**0.046**	**0.001**	**1.34**	**10**	**0.201**	**0.001**	**10.42**	**10**	**0.000**	**0.008**
GDA	**15.41**	**1**	**0.000**	**0.001**	**4.93**	**1**	**0.026**	**0.000**	**4.87**	**1**	**0.027**	**0.000**
GDA*Reference amount	**31.78**	**2**	**0.000**	**0.005**	**8.05**	**2**	**0.000**	**0.001**	**25.85**	**2**	**0.000**	**0.004**
GDA*Country	0.60	5	0.702	0.000	2.05	5	0.069	0.001	**27.22**	**5**	**0.000**	**0.010**
GDA*Reference amount*Country	**2.25**	**10**	**0.013**	**0.002**	1.19	10	0.292	0.001	**13.42**	**10**	**0.000**	**0.010**
GDA*Healthfulness	**119.96**	**1**	**0.000**	**0.009**	**17.13**	**1**	**0.000**	**0.001**	**5.88**	**1**	**0.015**	**0.000**
GDA*Healthfulness*Reference amount	**3.25**	**2**	**0.039**	**0.000**	2.84	2	0.059	0.000	**6.83**	**2**	**0.001**	**0.001**
GDA*Healthfulness*Country	**3.44**	**5**	**0.004**	**0.001**	1.24	5	0.288	0.000	**10.26**	**5**	**0.000**	**0.004**
GDA*Healthfulness*Reference amount*Country	**3.07**	**10**	**0.001**	**0.002**	1.05	10	0.401	0.001	**17.36**	**10**	**0.000**	**0.013**

Bold print indicates significant results at $p \leq 0.05$. Greenhouse-Geisser corrections were utilised to correct for the violation of the sphericity assumption where appropriate.



*The marked area indicates the objective healthfulness of the two products, as calculated via the SSAg/1 classification scheme. The upper line marks the less healthful product and the bottom line marks the more healthful product.

Biscuits (objective healthfulness ratings: more healthful = 5; less healthful = 12) when presented in varying reference amounts (half typical=9g; typical=18g; per 100g).
Sandwiches (objective healthfulness ratings: more healthful=0; less healthful=4) when presented in varying reference amounts (half typical=125g; typical=250g; per 100g).
Yogurts (objective healthfulness ratings: more healthful=0; less healthful=3) when presented in varying reference amounts (half typical=75g; typical=150g; per 100g).

Figure 1. Participants mean rescaled healthfulness ratings for more and less healthful products versus objective healthfulness scores, across three different categories.

DISCUSSION

Overall, participants correctly ranked foods according to their objective healthfulness as defined by their risk nutrients alone and could distinguish between more and less healthful variants of foods. This was the case both when seeing the nutrition label in absolute values and with additional GDA labelling. Over and above being presented with reference amount-based (absolute) values for energy and the four nutrients (fat, saturated fat, sugars, salt), product healthfulness evaluations did not change much by adding percentages of proposed daily reference quantities (GDAs) to the label. This is in line with previous research.[23–25]

General healthfulness associations with the three product categories do not appear to have had a strong influence on product ratings. Additionally, country-wise analyses revealed differences in the overall evaluation of more versus less healthful products. Comparing country ratings across all three foods, French consumers displayed the tendency to rate most extreme, that is, less healthful variants the lowest and more healthful variants the highest. Opposite results were obtained for Poland where product ratings were most similar and respondents did not differentiate as much. Explanations for this may include the role of foods in each country, the relevance of health and nutrition and the history of nutrition labelling, that is, presence on the market, exposure to and familiarity with this type of product information and media debate.[26]

Selection of appropriate reference amounts, for example, 'per 100 g', 'per portion' or 'per unit', is central to current debates within the scientific community regarding nutrition profiling systems and their ability to classify foods according to their healthfulness.[27–30] These results show that across the three food categories, consumers do factor the reference amount, that is, the quantity of food for which the nutritional information is being presented, into their judgements of healthfulness.

Therefore, appropriate reference amounts are also of importance for the effective presentation of nutritional information.[31] Previous research has shown that there is a tendency for increased consumption of foods perceived to be more healthful,[31–33] and that portion size can have a considerable effect on energy intake.[34–38] This calls for further research, in conjunction with other attributes of a nutrition label, for example, format, wording, design and possibly the interaction with other cues, for example, the perceived visual size of a portion on a plate or the number of units in a pack, the aim being to better understand what determines consumers' understanding of the information provided and how they translate this into food choice and consumption.

This study is confined to healthfulness evaluations of product descriptions with no cues other than a brief product name, which was the same for the more and less healthful products, and information about energy and the four risk nutrients the current legislation[9] requires. One could argue that communication of risk nutrients alone does not accurately communicate healthfulness; an example being between wholegrain and white bread where the risk nutrients could be the same but the presence of fibre and other micronutrients render the former as more healthful. This is an issue that warrants future research which could explore the use of further cues, such as positive ingredients and their impact on perceptions of product healthfulness. For that purpose, alternative objective health scoring systems would need to be considered.

CONFLICT OF INTEREST

MMR and CH's research centre has provided consultancy to and received travel funds to present research results from organisations supported by food and drinks companies.

AUTHOR CONTRIBUTIONS

All authors played an important role in interpreting the results and drafting the manuscript. The study was conceived by MMR, CH, JK and JW. MMR, CJ and CH analysed the data, SH and MMR conceived the design of the paper, SH drafted the manuscript and all authors commented on and approved the final version.

REFERENCES

1 Grunert KG, Fernández-Celemín L, Wills JM, Storcksdieck S, Nureeva L. Use and understanding of nutrition information on food labels in six European countries. J Public Health 2010; 18: 261–277.

2 Kelly B, Hughes C, Chapman K, Chun-Yu Louie J, Dixon H, Crawford J et al. Consumer testing of the acceptability and effectiveness of front-of-pack food labelling systems for the Australian grocery market. Health Promot Int 2009; 24: 120–129.

3 Borgmeier I, Westenhoefer J. Impact of different food label formats on healthiness evaluation and food choice of consumers: a randomized-controlled study. BMC Public Health 2009; 9: 184–195.

4 Feunekes GIJ, Gortemaker IA, Willems AA, Lion R, Van den Kommer M. Front-of-pack nutrition labelling: testing effectiveness of different nutrition labelling formats front-of-pack in four European countries. Appetite 2008; 50: 57–70.

5 Van Kleef E, Van Trijp H, Paeps F, Fernández-Celemín L. Consumer preferences for front-of-pack calories labelling. Public Health Nutr 2008; 11: 203–213.

6 Antonuk B, Block LG. The effect of single serving versus entire package nutrition information on consumption norms and actual consumption of a snack food. J Nutr Educ Behav 2006; 38: 365–370.

7 Sanders T, Breakfast Cereals Up Front – Debating the Issue of Food Labelling. Presentation to the Breakfast Cereal Information Service, London, 7 March 2006. Available at: http://www.powershow.com/view/128a54-ZmI2Z/Breakfast_Cereals_ Up_Front_Debating_the_Issue_of_Food_Labelling_powerpoint_ppt_presentation. Accessed 11 September 2013.

8 Lobstein T, Davies S. Defining and labelling 'healthy' and 'unhealthy' food. Public Health Nutri 2008; 12: 331–340.

9 EU Regulation No 1169/2011 on the provision of food information to consumers. Europa, Brussels. Available at: http://eurlex.europa.eu/LexUriServ/LexUriServ.do? uri = OJ:L:2011:304:0018:0063:EN:PDF. Accessed 11 September 2013.

10 Cowburn G, Stockley L. Consumer understanding and use of nutrition labelling: a systematic review. Public Health Nutr 2005; 8: 21–28.

11 Temple N, Fraser J. Food labels: a critical assessment. Nutrition 2014; 30: 257–260.

12 Faulkner GP, Pourshahidi LK, Wallace JMW, Kerr MA, McCrorie TA, Livingstone MB. Serving size guidance for consumers: is it effective? Proc Nutr Soc 2012; 71: 610–621.

13 Scarborough P, Rayner M, Stockley L. Developing nutrient profile models: a systematic approach. Public Health Nutr 2006; 10: 330–336.

14 Drewnowski A, Fulgoni V. Nutrient profiling of foods: creating a nutrient-rich food index. Nutr Rev 2008; 66: 23–39.

15 Visschers VHM, Siegrist M. Applying the evaluability principle to nutrition table information. How reference information changes people's perception of food products. Appetite 2009; 52: 505–512.

16 Provencher V, Polivy J, Herman CP. Perceived healthiness of food. If it's healthy, you can eat more!. Appetite 2009; 52: 340–344.

17 Viswanathan M. The influence of summary information on the usage of nutrition information. J Public Policy Marketing 1994; 13: 48–60.

18 Viswanathan M, Hastak M. The role of summary information in facilitating consumers' comprehension of nutrition information. J Public Policy Marketing 2002; 21: 305–318.

19 Carels RA, Konrad K, Harper J. Individual differences in food perceptions and calorie estimation. An examination of dieting status, weight, and gender. Appetite 2007; 49: 450–458.

20 Wiseman M. The COMA Report: dietary reference values for food energy and nutrients for the United Kingdom. British Food Journal 1992; 94: 7–9.

21 Storcksdieck S, Fernández-Celemín L, Larrañaga A, Egger S, Wills JM, Hodgkins C et al. Penetration of nutrition information on food labels across the EU-27 plus Turkey. Eur J Clin Nutr 2010; 64: 1379–1385.

22 National Readership Survey (2007) Lifestyle data. Available at: http://www.nrs.co. uk/lifestyle-data/. Accessed 17 December 2013.

23 Rayner M, Scarborough P, Stockley L, Nutrient profiles. Options for definitions for use in relation to food promotion and children's diet. British Heart Foundation Health Promotion Research Group, Department of Public Health, University of Oxford; 2004. Available at: http://www.food.gov.uk/multimedia/pdfs/nutrientpro filingfullreport.pdf. Accessed 17 December 2013.

24 Malam S, Clegg S, Kirwan S, McGinigal S, Raats MM, Shepherd R et al. Comprehension and use of UK nutrition signpost labelling schemes. London, Food Standard Agency; 2009. Available at: http://www.food.gov.uk/multimedia/pdfs/ pmpreport.pdf. Accessed 17 December 2013.

25 Grunert KG, Wills JM, Fernández-Celemin L. Nutrition knowledge, and use and understanding of nutrition information on food labels among consumers in the UK. Appetite 2010; 55: 177–189.

26 Aschemann-Witzel J, Grunert KG, van Trijp H, Bialkova S, Raats MM, Hodgkins C et al. Effects of nutrition label format and product assortment on the healthfulness of food choice. Appetite 2013; 71: 63–74.

27 Van Herpen E, Seiss E, Van Trijp HCM. The role of familiarity in front-of-pack label evaluation and use: A comparison between the United Kingdom and The Netherlands. Food Quality and Preference 2012; 26: 22–34.

28 Drewnowski A, Maillot M, Darmon N. Should nutrient profiles be based on 100 g, 100 kcal or serving size? Eur J Clin Nutr 2009; 63: 898–904.

29 Scarborough P, Arambepola C, Kaur A, Bhatnagar P, Rayner M. Should nutrient profile models be 'category specific' or 'across-the-board'? A comparison of the two systems using diets of British adults. Eur J Clin Nutr 2010; 64: 553–560.

30 EFSA. The setting of nutrient profiles for foods bearing nutrition and health claims pursuant to Article 4 of the Regulation (EC) N1 1924/2006- Scientific opinion of the panel on dietetic products, nutrition and allergies. EFSA J 2008; 644: 1–44.

31 Azais-Braesco V, Goffi C, Labouze E. Nutrient profiling: comparison and critical analysis of existing systems. Public Health Nutr 2006; 9: 613–622.

32 Wansink B, Chandon P. Can "low-fat" nutrition labels lead to obesity? J Marketing Res 2006; 43: 605–617.

33 Faulkner GP, Pourshahidi LK, Wallace JMW, Kerr MA, McCaffrey TA, Livingstone MBE. Perceived 'healthiness' of foods can influence consumers' estimations of energy density and appropriate portion size. Int J Obes 2014; 38: 106–112.

34 Steenhuis I, Vermeer WM. Portion size: review and framework for interventions. Int J Behav Nutr Phys Act 2009; 6: 1–10.

35 Dilberti N, Bordi PL, Conklin MT, Roe LS, Rolls BJ. Increased portion size leads to increased energy intake in a restaurant meal. Obes Res 2004; 12: 562–568.

36 Rolls BJ, Roe JS, Meengs JS. Larger portion sizes lead to sustained increases in energy intake over 2 days. J Am Diet Assoc 2006; 106: 543–549.

37 Flood JE, Roe LS, Rolls BJ. The effect of increased beverage portion size on energy intake at a meal. J Am Diet Assoc 2006; 106: 1984–1990.

38 Rozin P, Kabnick K, Pete E, Fischler C, Shields C. The ecology of eating: Smaller portion sizes in France than in the United States help explain the French paradox. Psychol Sci 2003; 14: 450–454.